MATHÉMATIQUES ET SCIENCES HUMAINES

BERNARD MONJARDET

Sur diverses formes de la « règle de Condorcet » d'agrégation des préférences

Mathématiques et sciences humaines, tome 111 (1990), p. 61-71
<http://www.numdam.org/item?id=MSH_1990__111__61_0>

Math. Inf. Sci. hum., (28e année, n°111, *1990*, p.61-71)

SUR DIVERSES FORMES DE LA "RÈGLE DE CONDORCET" D'AGRÉGATION DES PRÉFÉRENCES

Bernard MONJARDET[1]

RÉSUMÉ - *Nous appelons ici règle ou procédure de Condorcet la procédure d'agrégation d'ordres des préférences individuelles en un ordre collectif consistant à chercher un ordre recueillant le nombre maximum de suffrages sur toutes les préférences par paires qu'il exprime. La définition précise de cette procédure et la raison de son appellation se trouvent dans l'introduction. Le reste du texte présente de multiples formes équivalentes pour la définir et donne des indications historiques et bibliographiques sur ses redécouvertes ultérieures.*

ABSTRACT - On several forms of the "Condorcet's rule" of aggregation of preferences.
This note bears on a rule of aggregation of individual preferences into a collective preference called here "the Condorcet's rule" and often called in the litterature of social choice "the Kemeny's rule" or the "median procedure". First we give a definition of this rule and why it is now attributed to Condorcet. Then we give about twenty five equivalent definitions of this rule and some historical and bibliographical comments on its many subsequent rediscoveries.

1. INTRODUCTION

On considère une assemblée $V = \{1,....r, s,....v\}$ de v "votants" ; chaque votant exprime ses préférences sur un ensemble $X = \{1,...i, j, k, ...n\}$ de n"objets" (candidats, options, etc...). On suppose que cette préférence est toujours un *ordre strictement total* sur X ; autrement dit, c'est une relation O sur X, *transitive* (i O j et j O k impliquent i O k), *asymétrique* (i O j implique j O^c i où O^c désigne la relation complémentaire de O, i.e. $O^c = X^2 - O$), et *strictement totale* ($i \neq j$ et i O^c j impliquent j O i). Comme dans la suite nous ne considérons que des ordres strictement totaux, nous les appellerons simplement des ordres ; noter qu'ils sont antiréflexifs (pour tout i, i O^c i). On note $\mathcal{O}_X$ ou simplement $\mathcal{O}$ l'ensemble des n! ordres définis sur X ($|X| = n$).

Si i et j sont deux éléments distincts de X et que pour le votant r de V, on ait i O_r j, on dit que r préfère i à j. On pose :

$$v_{ij} = |\{r \in V : i\, O_r\, j\}|$$

v_{ij} est donc le nombre de votants de l'assemblée préférant i à j.

---

[1] Université Paris 5 et CAMS, 54 boulevard Raspail 75270 Paris Cedex 06

On considère le problème d'optimisation (C) suivant, ou "problème de Condorcet" (cf. plus loin la raison de cette appellation) :

Chercher un ordre C de $\mathcal{O}$ tel que

$$(C) \qquad \sum_{iCj} v_{ij} = \max_{O \in \mathcal{O}} \Big[\sum_{iOj} v_{ij}\Big]$$

Pour un ordre O donné la quantité $\sum_{iOj} v_{ij}$ est la somme pour chaque couple - i préféré à j - de l'ordre O, des nombres de votants qui préfèrent i à j ; c'est donc une mesure de l'"accord" de l'assemblée si O est retenu comme ordre collectif, agrégation des ordres individuels des votants. Il est alors naturel de choisir comme ordre collectif un ordre maximisant cet accord, c'est-à-dire un ordre solution du problème (C). Un tel ordre sera aussi appelé, pour des raisons explicitées ultérieurement, *ordre médian* des ordres $O_1,...O_r,...O_v$.

Il existe un cas où la solution du problème (C) est très simple. Considérons la relation majoritaire (stricte) M associée aux préférences des votants : i M j si et seulement si $v_{ij} > v_{ji}$. Si cette relation majoritaire est sans circuit, c'est-à-dire, si par exemple, il n'existe pas i, j, k avec i M j, j M k et k M i, elle est contenue dans un ordre partiel strict (obtenu par "fermeture transitive" de M), et lui même prolongeable, de différentes façons, en un ordre (strictement total) ; les ordres médians sont alors tous les ordres obtenus de cette manière. En particulier si le nombre de votants est impair et la relation majoritaire transitive, elle constitue l'unique ordre solution du problème (C).

On sait que la définition de cette relation majoritaire remonte à Condorcet, ainsi que la mise en évidence du fait que cette relation peut présenter des circuits, ou en termes de préférences des intransitivités. Dans le texte de 1952, où en particulier Guilbaud montrait l'importance du travail de Condorcet, il dénommait "effet Condorcet" ce phénomène. Lorsqu'il y a effet Condorcet la relation majoritaire ne peut plus définir (directement ou par complétion ordinale) la préférence collective, si tant est qu'on exige de cette dernière d'être un ordre, condition naturelle dans les problèmes de décision. Il faut donc trouver un autre moyen d'agréger les ordres individuels en un ordre collectif. Et Guilbaud ajoutait (p. 50) : "Condorcet ne peut se résigner à conclure qu'on ne peut attribuer aucune opinion cohérente au corps électoral...... Il cherche un moindre mal, c'est-à-dire parmi toutes les opinions *cohérentes* , celle qui est appuyée par le plus grand nombre possible de suffrages". Comme ici "opinion cohérente" signifie ordre (total), on voit que Guilbaud attribue à Condorcet la règle définie par le problème d'optimisation (C) : maximiser la somme des suffrages v_{ij} sur toutes les préférences i O j d'un ordre O. On sait qu'on a pu en fait hésiter à attribuer cette règle à Condorcet lui-même ; s'il existe dans l'Essai un ou deux passages, correspondant à la citation de Guilbaud, où Condorcet semble clairement désigner cette méthode pour résoudre le problème de l'"effet Condorcet", il en est d'autres nettement moins clairs ; d'autre part les méthodes explicites de calcul qu'il donne pour obtenir l'ordre collectif ne conduisent pas nécessairement à une solution de (C), lorsque v est supérieur à 3 (ce n'est guère étonnant, compte tenu de la difficulté que peut présenter ce calcul). C'est donc plus à partir d'une compréhension globable de la démarche de Condorcet, notamment dans ses aspects probabilistes, qu'on est fondé à croire que Condorcet est l'inventeur de la procédure définissant ce que nous appelons les ordres médians (cf. à ce sujet Young, 1988 ainsi que ce numéro de *Mathématiques, Informatique et Sciences humaines* et Crépel, 1990)[2]. Dans la suite les

[2] Nous n'oublions pas les contributions originales de Michaud (1982, 1985) qui se proclame ardent défenseur de Condorcet. Tout en basant son argumentation sur la phrase de Guilbaud citée dans l'introduction, il estime que cet auteur a "ignoré le sens profond" de l'effet Condorcet, et que plus généralement avec beaucoup d'autres (Kendall, Arrow, etc..) il a cru en l'impossibilité d'une règle d'agrégation acceptable, règle justement fournie par Condorcet. (Nous lui laissons bien sûr la responsabilité de ces affirmations, qui entre autres nous semblent révéler une

termes *règle* ou *procédure de Condorcet* désigneront donc toujours l'association à un profil d'ordres individuels des ordres médians, *i.e.* des ordres solutions du problème (C). Redisons que cette règle coïncide exactement avec la règle majoritaire (dite aussi communément règle de Condorcet) lorsque celle-ci donne un ordre (strictement total), et que dans les autres cas elle peut conduire à plusieurs ordres solutions (phénomène qui ne semble pas avoir été remarqué par Condorcet).

L'objet principal de ce texte est de donner une liste de problèmes d'optimisation sur l'ensemble $\mathcal{O}_X$ des ordres qui, étant équivalents au problème de Condorcet, définissent la même règle d'agrégation. Comme d'une part la formulation "moderne" du problème (C) est assez récente et que d'autre part il est très simple mais non toujours évident de montrer que les différents problèmes d'optimisation considérés sont équivalents, ceci explique que la procédure de Condorcet ait été retrouvée plusieurs fois sous des appellations variées et de façon indépendante, la "règle de Kemeny" étant l'exemple le plus connu de ces redécouvertes.

Dans la section suivante, nous donnons les définitions et notations nécessaires. En particulier nous introduisons deux codages vectoriels des ordres. Dans la troisième section, nous donnons la liste des problèmes d'optimisation équivalents au problème de Condorcet et la preuve de ces équivalences. La dernière section comporte des indications historiques et bibliographiques sur les problèmes présentés.

2. DEFINITIONS ET NOTATIONS

$X = \{1,...i, j, ...n\}$; $V = \{1, ...r, s,...v\}$

$\mathcal{O} = \{$n! ordres (strictement totaux) définis sur X$\}$

$\Pi = (O_1,...O_r,...O_v)$, où $O_r \in \mathcal{O}$ pour tout $r \in V$, est appellé un *profil* d'ordres

$\mathcal{O}^v = \{(n!)^v$ profils $\Pi\}$

Pour i, j distincts, éléments de X, on pose :

$$v_{ij} = |\{r \in V : i\, O_r\, j\}| \quad ; \quad \text{on a} \quad v_{ij} + v_{ji} = v$$
$$w_{ij} = v_{ij} - v_{ji} = 2v_{ij} - v = v - 2v_{ji}$$
$$w^+_{ij} = w_{ij} - \min(v_{ij}, v_{ji}) = \begin{cases} w_{ij} & \text{si } w_{ij} \geq 0 \\ 0 & \text{si } w_{ij} \leq 0 \end{cases}$$

Pour deux ordres O et O' de $\mathcal{O}$, on pose :

$$d_\Delta(O,O') = |O \Delta O'| = |O - O'| + |O' - O|$$
$$d_K(O,O') = \frac{1}{2} d_\Delta(O,O') = |O - O'| = |O' - O|$$
$$\tau(O,O') = 1 - \frac{4d_K(O,O')}{n(n-1)}$$

d_Δ (O,O') est la classique distance de la différence symétrique entre deux relations ; sa moitié d_K(O,O') est le nombre de (paires en) *désaccords* entre O et O', i.e. le nombre de paires $\{i,j\}$ d'objets pour lesquelles i est préféré à j dans un ordre et j préféré à i dans l'autre ; elle varie donc entre 0 et $\binom{n}{2}$. Normalisée pour varier de + 1 en cas d'absence de désaccord entre O et O', à -1 en

certaine incompréhension de l'analyse subtile que Guilbaud fait du théorème d'Arrow).

cas de nombre maximum de désaccords entre O et O', elle donne le classique tau de Kendall τ (O,O') entre les ordres O et O'. Il est aussi bien connu que $d_K(O,O')$ est identique à la distance des "transpositions consécutives" entre deux ordres (totaux), c'est-à-dire au nombre minimum d'inversions de deux objets consécutifs, permettant de passer du premier ordre au second.

Pour un profil $\Pi = (O_1,...O_r,O_s, ...O_v)$, on pose :

$$K(\Pi) = \sum_{(r,s)\in V^2} \frac{\tau(O_r,O_s)}{v^2} \qquad\qquad U(\Pi) = 2 \sum_{\{r,s\}\subseteq V} \frac{\tau(O_r,O_s)}{v(v-1)}$$

$K(\Pi)$ est la moyenne des tau de Kendall prise sur tous les couples d'ordres du profil Π ; $U(\Pi)$ est cette même moyenne, mais prise sur toutes les paires d'ordres. K et U sont deux coefficients de concordance entre les ordres du profil reliés par la formule (évidente) :

$$U = \frac{v\,K - 1}{v - 1}$$

U s'appelle aussi le coefficient de Kendall-Ehrenberg du profil Π.

Nous définissons maintenant deux codages associant à tout ordre de $\mathcal{O}$ un vecteur de l'espace euclidien $\mathbb{R}^{n(n-1)}$:

- le *codage caractéristique* : $O \to x^O$

Il est défini en posant

$$x^O{}_{ij} = 1 \Leftrightarrow i\,O\,j \qquad\qquad x^O{}_{ij} = 0 \Leftrightarrow j\,O\,i$$

La suite x^O des valeurs $x^O{}_{ij}$ (ordonnée suivant un ordre arbitraire mais fixé des couples) est donc un vecteur de $\mathbb{R}^{n(n-1)}$, à $\frac{n(n-1)}{2}$ composantes égales à 1, et $\frac{n(n-1)}{2}$ composantes nulles ; on l'appelle *vecteur caractéristique* de O. Si $\Pi = (O_1,...O_r,...O_v)$ est un profil, on note x_r le vecteur associé à O_r et $X(\Pi)$, ou simplement X, la suite $(x_1,...,x_r,...x_v)$ des vecteurs caractéristiques des x_r ; on note $\bar{x}$ (Π), ou simplement $\bar{x}$, le vecteur moyen $\bar{x} = \sum_{r=1}^{v} \frac{x_r}{v}$; ses coordonnées sont les n(n-1) nombres v_{ij} $(i\neq j)$.

- le *codage caractéristique centré* : $O \to y^O$

Il est défini en posant $y^O{}_{ij} = x^O{}_{ij} - x^O{}_{ij}$, soit

$$y^O{}_{ij} = 1 \Leftrightarrow i\,O\,j \qquad\qquad y^O{}_{ij} = -1 \Leftrightarrow j\,O\,i$$

Le vecteur y^O de $\mathbb{R}^{n(n-1)}$ ainsi défini est appelé le *vecteur caractéristique centré* associé à l'ordre O. On note $Y(\Pi)$ ou simplement Y, la suite $(y_1,...y_r,...y_v)$ des vecteurs centrés associés aux ordres d'un profil $\Pi = (O_1,...O_r,...O_v)$; on note $\bar{y}(\Pi)$, ou simplement $\bar{y}$ le vecteur moyen :

$\bar{y} = \sum_{r=1}^{v} \frac{y_r}{v}$; ses coordonnées sont les n(n-1) nombres w_{ij} $(i \neq j)$.

Dans l'espace euclidien $\mathbb{R}^{n(n-1)}$, d_E dénote la distance euclidienne, $||t||$ la norme du vecteur t, $\langle t,t' \rangle$ le produit scalaire des deux vecteurs t et t' ; on note d_1 la distance de la norme L_1 entre deux vecteurs, i.e. la somme des valeurs absolues de leurs différences de coordonnées. Nous donnons sans démonstrations (elles sont toutes élémentaires) quelques relations liant les quantités introduites ci-dessus.

(a) $\sum_{iOj} v_{ij} + \sum_{jOi} v_{ji} = v\frac{n(n-1)}{2}$

(b) $\sum_{iOj} v_{ji} = \sum_{r=1}^{v} d_k(O_r,O)$

(c) $d_\Delta(O,O') = 2d_K(O,O') = \frac{1}{2}\, d_1(y,y') = \frac{1}{4}\, d_E^2\,(y,y') = d_E^2\,(x,x')$

(d) $\tau(O,O') = \cos(y,y') = \frac{\langle y,y' \rangle}{n(n-1)} = 1 - \frac{d_E^2(y,y')}{2n(n-1)}$

(e) $\sum_{r=1}^{v} \cos(y_r,y) = \frac{v\,\langle \bar{y},y \rangle}{n(n-1)} = \frac{v||y||}{[n(n-1)]^{\frac{1}{2}}} \cos(\bar{y},y)$

3. PROBLÈMES D'OPTIMISATION ÉQUIVALENTS AU PROBLÈME DE CONDORCET.

Tous les problèmes d'optimisation sur $\mathcal{O}$ que nous allons lister ci-dessous sont équivalents au problème de Condorcet. Comme celui-ci, ils se présentent sous la forme suivante :

Maximiser (ou Minimiser) sur l'ensemble $\mathcal{O}$ des ordres une fonction dépendant du profil $\Pi = (O_1,...O_r,..O_v)$ donné et de l'ordre O considéré ; la fonction fait intervenir les différentes quantités ou coefficients présentés dans la section précédente. Ci-dessous nous notons simplement par x (respectivement y) le vecteur caractéristique (resp. le vecteur caractéristique centré) associé à un ordre O arbitraire ; ses coordonnées sont donc notées x_{ij} (respectivement y_{ij}). Certains problèmes sont clairement équivalents ; dans ce cas ils ont été mis sur la même ligne.

Nous commençons par la liste des problèmes de maximisation, donc de la forme :

$$\mathbf{MAX}\ [f(\Pi,O)\ ,\ O \in \mathcal{O}]$$

avec $f(\Pi,O)$ égal à :

(1) $\sum_{iOj} v_{ij}$ (1bis) $\sum_{i \neq j} v_{ij}\ x_{ij}$

(2) $\sum_{iOj} w_{ij}$ (2bis) $\sum_{i \neq j} w_{ij}\ x_{ij}$

(3) $\sum_{iOj} w^{+}_{ij}$ (3bis) $\sum_{i \neq j} w^{+}_{ij}\ x_{ij}$

(4) $\sum_{r=1}^{v} \tau(O_r, O)$ (4bis) $\sum_{r=1}^{v} \cos(y_r, y)$

(5) $K(\Pi + O)$ (5bis) $U(\Pi + O)$

(6) $< \bar{y}, y >$ (6bis) $\cos(\bar{y}, y)$

Le problème (1) n'est autre que le problème de Condorcet de l'introduction ; en faisant intervenir le codage caractéristique de l'ordre O, il s'écrit sous la forme (1bis), qui a l'avantage de faire apparaître son caractère linéaire. Il en est de même lorsqu'on écrit (2) ou (3) sous les formes (2bis) ou (3bis). L'équivalence de (1), (2) et (3) résulte des relations $w_{ij} = 2 v_{ij} - v$ et $w^{+}_{ij} = w_{ij} - \min (v_{ij}, v_{ji})$. Les fonctions (4) et (4bis) sont les mêmes d'après la relation (d). L'équivalence des problèmes (4) et (5) provient de l'égalité $(v + 1)^2 K(\Pi + O) = v^2 K(\Pi) + 2\sum_{r=1}^{v} \tau(O_r, O) + 1$; celle des problèmes (5) et (5bis) de la relation $U = \frac{vK - 1}{v-1}$. La relation (e) montre l'équivalence des problèmes (4bis), (6) et (6bis). Reste à montrer par exemple l'équivalence de (1) et (4) ce qui sera fait plus loin, en utilisant leur équivalence commune à un problème de minimisation.

Nous donnons maintenant la liste des problèmes de minimisation :

$$\mathbf{MIN}\ [f(\Pi, O), O \in \mathcal{O}]$$

avec $f(\Pi, O)$ égal à :

(7) $\sum_{iOj} v_{ji}$ (7bis) $\sum_{i \neq j} v_{ji}\, x_{ij}$

(8) $\sum_{iOj} w_{ji}$ (8bis) $\sum_{i \neq j} w_{ji}\, x_{ij}$

(9) $\sum_{iOj} w^{+}_{ji}$ (9bis) $\sum_{i \neq j} w^{+}_{ji}\, x_{ij}$

(10) $\sum_{r=1}^{v} d_{\Delta}(O_r, O)$ (10bis) $\sum_{r=1}^{v} d_k(O_r, O)$ (10.3) $\sum_{r=1}^{v} d_1(y_r, y)$ (10.4) $\sum_{r=1}^{v} d^2_E(y_r, y)$ (10.5) $\sum_{r=1}^{v} d^2_E(x_r, x)$

(11) $d_E(\bar{x}, x)$ (11 bis) $d_E(\bar{y}, y)$

L'équivalence des problèmes (7), (7bis), (8), (8bis), (9), (9bis) se montre comme pour les problèmes (1), ...(3bis). L'équivalence de (1) et (7) est immédiate d'après la relation (a) : de même pour (7) et les différents problèmes (10), à partir des relations (c) et (d). D'autre part dans l'espace euclidien $\mathbb{R}^{n(n-1)}$, la formule de Huyghens permet d'écrire :

$$\sum_{r=1}^{v} d_E^2\,(y_r,y) = \sum_{r=1}^{v} d_E^2\,(y_r,\bar{y}) + v\; d_E^2\,(\bar{y},y)$$

on en déduit l'équivalence des problèmes (10.4) et (11bis) ; de même pour (10.5) et (11). Enfin le problème de minimisation (10.4) est équivalent au problème de maximisation (4) du fait de la relation (d).

On peut remarquer que les codages caractéristiques des ordres utilisés ci-dessus sont "sphériques" en ce sens que la norme des vecteurs correspondants est constante. La plupart des équivalences montrées *supra* sont valables dans le cadre géométrique général d'un nuage de points sur une sphère. En particulier elles sont valables lorsqu'on prend un autre codage classique des ordres, celui défini dans $\mathbb{R}^n$ par les "rangs" ou les "scores" des objets de X ; on obtient ainsi différentes caractérisations de la classique procédure de Borda, basée sur les sommes des rangs, dont on reparlera plus loin (cf. à ce sujet, Monjardet, 1985).

Pour terminer cette section, notons que nous avons choisi une écriture relationnelle ou vectorielle des problèmes d'optimisation ci-dessus ; on aurait pu aussi prendre une écriture matricielle, qu'on trouvera chez certains auteurs. Nous signalerons aussi plus loin une présentation "permutationnelle" de certains de ces problèmes, qui peut toutefois être assez lourde à écrire. Enfin nous ne prétendons pas avoir épuisé toutes les formes d'écriture du problème ; on pourrait par exemple l'écrire en termes de théorie des graphes ; nous avons aussi omis une forme duale intéressante des problèmes (7bis) et (8bis) (Merchant et Rao, 1977).

4. INDICATIONS HISTORIQUES ET BIBLIOGRAPHIQUES.

Faisons d'abord une remarque ; nous n'avons considéré ci-dessus que le cas étudié par Condorcet où les préférences individuelles et collectives des votants sont des ordres (strictement totaux). Dans les textes dont nous allons parler, on considère souvent des cas plus généraux, notamment celui où ces préférences sont des préordres totaux (l'indifférence -transitive - est admise). La procédure Condorcet peut bien sûr être généralisée, de différents points de vue, à ces cas. Mais pour ne pas alourdir cet exposé et rester dans le cadre qui était celui de Condorcet, nous ne considérons dans les exposés généraux que ce qui concerne le cas particulier des ordres.

On peut apparemment regrouper les différentes redécouvertes de la procédure Condorcet autour de trois thèmes : la notion de médiane dans un espace métrique (Kemeny, 1959, etc...) ; l'idée de minimisation des désaccords ou de maximisation des accords (Brunk, 1960 à partir de Kendall, 1955, etc...) ; l'utilisation du tau de Kendall (Hays, 1960, etc...).

Le fait que l'on puisse considérer la relation majoritaire définie par Condorcet comme une médiane dans un certain espace apparaît dès l'article de Guilbaud de 1952 ; c'est toutefois l'aspect algébrique de la médiane, plus que son aspect métrique, qui est développé (pour la relation entre ces deux aspects, cf. Monjardet 1990). Inversement, Kemeny en 1959, (cf. aussi Kemeny et Snell, 1962) pose d'emblée le problème de la recherche de l'ordre collectif comme celui de définir un espace métrique sur l'ensemble des ordres, puis d'y rechercher l'analogue des médianes ou des moyennes en statistique ; par contre le lien avec la procédure majoritaire n'apparaît pas. Kemeny part d'une liste d'axiomes qu'on peut vouloir imposer à une distance entre ordres et montre que cette distance est alors celle de la différence symétrique $d_\Delta(O,O')$, qu'il écrit sous la forme $\frac{1}{2}$ d_1 (y,y'). Il propose ensuite de rechercher l'ordre le "plus proche" des ordres donnés, cette proximité étant définie soit au sens de la somme des distances d'un ordre considéré aux ordres donnés, soit au sens de la somme des carrés de ces distances ; le premier choix est considéré comme la recherche d'"ordres médians" (en fait, ce terme n'est pas dans Kemeny, mais dans les travaux français décrits ci-dessous) ; le second comme la recherche d'"ordres moyens". Le premier choix qui conduit au problème d'optimisation (10.3) correspond donc à la procédure Condorcet ; Kemeny note qu'elle peut conduire à plusieurs solutions. Cette "Kemeny rule", bien qu'assez connue aux Etats-Unis, grâce au livre de Kemeny et Snell, n'y suscitera des développements qu'à partir des années soixante-dix. Indépendamment, en France, l'usage et les propriétés de la médiane métrique comme résumé de données ordinales avaient été étudiés par Barbut (1961, 1966, 1967) puis différents auteurs (notamment Jacquet-Lagrèze, 1969, 1971, Monjardet, 1973 ; cf. aussi Barbut et Monjardet, 1970, et Degenne, 1972) : ils avaient en particulier considéré le problème de la recherche des ordres médians (sous la forme des problèmes (10)), et noté la relation simple mais fondamentale $\sum_{r=1}^{v} d_K(O_r,O) = \sum_{iOj} v_{ji}$. De cette relation découle notamment le fait mentionné dans l'introduction (et redécouvert maintes fois) que lorsque la relation majoritaire est sans circuit, on peut en déduire les ordres médians (par complétion) ; en particulier, si la relation majoritaire est un ordre c'est l'unique ordre médian. Notons à ce propos que d'autres règles que la procédure Condorcet satisfont à cette condition d'"extension" de la règle majoritaire (cf. l'article de Bowman et Colantoni, 1973, cité ci-dessous, et dans le contexte un peu différent des fonctions de choix collectif, Fishburn, 1977). Des travaux sur les relations médianes du même type que les travaux français avaient été aussi mené en URSS en particulier grâce à Mirkin (1974). On peut aussi rattacher à ce courant métrique le travail de Bowman et Colantoni (1973) qui revient à considérer le problème de minimisation (11) avec une distance non nécessairement euclidienne et qui montre le caractère linéaire du problème (point déjà remarqué par De Cani, 1969, pour un problème plus général). Notons enfin que dans Levenglick (1974), on trouve l'équivalence entre les formulations 10.3 (règle de Kemeny explicitement citée) (6) et (11) de la procédure de Condorcet et que son travail conduira à une remarquable axiomatisation de cette procédure (Young et Levenglick, 1978).

Passons maintenant au thème "Maximisation des accords" ou "Minimisation des désaccords", correspondant aux deux problèmes équivalents :

Max $\sum_{iOj} v_{ij}$, *i.e.* le problème de Condorcet, et Min $\sum_{iOj} v_{ji}$; dans la deuxième formulation v_{ji} représente le désaccord "local" de l'assemblée si i est préféré à j dans l'ordre collectif, et il est naturel de chercher à minimiser le désaccord "global" obtenu par sommation des désaccords sur les couples de O. L'idée de minimiser des désaccords aussi bien pour trouver un ordre collectif que pour sélectionner un ou plusieurs vainqueurs apparaît dans un texte de Kendall (1955). Il écrit, par exemple (page 43) : "When we have to select a subset of the n objects as "elected" we shall in general, in the absence of complete unanimity, violate a number of preferences. Circumstances force us to do to some extent. The problem is to do so to the least possible extent". Cependant plusieurs de ses assertions (cf. notamment 7.f et 9, pages 48 et 49) semblent indiquer qu'il pensait que les "ordres de Borda" étaient les solutions optimales du problème. Rappelons qu'un ordre de Borda est obtenu à partir d'un profil, en classant les objets d'après l'ordre induit par les sommes des rangs qu'ils ont obtenus dans les différents ordres (sommes égales aux quantités $\sum_{i \neq j} v_{ji}$), et en départageant de manière arbitraire les *ex æquo* éventuels. Rappelons aussi que la procédure de Borda ne satisfait pas à la condition majoritaire de Condorcet, puisqu'elle ne donne pas nécessairement la relation majoritaire lorsque celle-ci est un ordre ; c'était d'ailleurs la principale critique qu'en faisait Condorcet. La raison de l'erreur de Kendall est toutefois intéressante. Il admet implicitement que la résolution du problème de

Condorcet (sous la forme (7)) peut se faire de façon séquentielle : on choisit comme premier élément de l'ordre cherché l'objet i créant le moins de désaccords $\sum_{j \neq i} v_{ji}$, donc l'objet de rang minimum ; puis parmi les objets restants, et sans recalculer les rangs, le second objet vérifiant le même critère, et ainsi de suite ; on aboutit bien ainsi, pratiquement par définition, à un ordre de Borda. Notons qu'il serait plus naturel de recalculer les rangs à chaque itération, en supprimant les préférences concernant les objets déjà classés : on aboutirait ainsi aux "ordres prudents" de Kohler, Arrow et Raynaud (1986). L'erreur de Kendall allait toutefois être corrigée par Brunk (1960), qui suivant " a suggestion of Kendall, that the number of expressed preferences which are violated be minimized" écrit exactement le problème de Condorcet - sous la forme (1) - et en montre l'équivalence avec la forme (7). La procédure Condorcet sous la forme (1), est encore redécouverte en 1973 par Blin et Whinston, qui l'appellent "The combinatorial optimization criterion" (cf. aussi 1975) ; le second de ces auteurs avec Adelsman (1977), s'aperçoit toutefois que leur méthode est équivalente à celle de Kemeny, ainsi qu'à une version duale du problème (8bis), qu'on trouve dans Merchant et Rao (1977). On notera que plusieurs de ces auteurs présentent le problème d'optimisation sous la forme de recherche d'une permutation de l'ensemble X, ce qui peut conduire à des formulations assez lourdes.

Venons en enfin au troisième thème, où la procédure Condorcet est redécouverte à partir de l'idée de "corrélation" entre ordres. En fait l'idée d'utiliser un coefficient de corrélation entre ordres pour obtenir un ordre consensus semble liée à Kendall (1942). Celui-ci parle d'ailleurs, comme Condorcet, d'estimation d'un "vrai" ordre ("true ranking") ; il utilise pour ce faire la méthode de Borda évoquée plus haut. Mais il montre qu'un ordre O ainsi obtenu maximise la quantité $\sum_{r=1}^{v} \rho(O_r, O)$, où ρ est le coefficient de corrélation de Spearman entre deux ordres. Si l'on remplace ce coefficient ρ par le coefficient τ, on aboutit au problème de maximisation (4), équivalent au problème de Condorcet. C'est exactement ce que propose de faire le statisticien Hays, dans un article paru la même année (1960) et dans la même revue (*Journal of the American Statistical Association*) que celui de Brunk. Dans le même ordre d'idées, au lieu de la formulation (4), on peut aussi utiliser les formulations (5) et (5bis), qui consistent à chercher l'ordre consensus comme celui qui ajouté aux ordres donnés maximise les deux coefficients de concordance entre ordres basés sur le tau, à savoir K et U (ceci est par exemple noté dans Marcotorchino et Michaud, 1979 et Barthélemy et Monjardet 1981).

Revenons à Hays pour signaler qu'il avait remarqué que résoudre ce problème (4) était équivalent à trouver l'ordre des lignes et des colonnes du tableau des n^2 coefficients v_{ij}, rendant maximum la somme des termes au-dessus de la diagonale. Ceci n'est rien d'autre qu'une formulation "permutationnelle" du problème (1). Elle est intéressante car elle montre que ce problème est un cas particulier d'un problème rencontré dans de nombreux contextes, par exemple dans celui de la "triangulation" d'un tableau d'échanges inter-industriels : ayant un tableau (carré) de coefficients c_{ij} ($1 \leq i$, $j \leq n$), trouver une permutation des lignes et des colonnes de ce tableau rendant maximum la somme des termes au dessus de la diagonale. On sait maintenant que ce problème général, comme le problème particulier de Condorcet, est NP complet, ce qui rend a posteriori quelque peu savoureuse la phrase de Hays, affirmant que la recherche d'un ordre médian "may be done quite simply by putting the rank-order data into a table showing the number of cases in which the row object is ranked above the column object and permuting rows and columns (simultaneously, so that rows and columns show the same order) until the sum of the entries above the diagonal is maximal".

Notons toutefois que malgré son caractère difficile (qui ne pouvait être connu en 1960) le problème de détermination effective des ordres médians a fait l'objet ces dernières années de progrès sensibles (*cf.* par exemple, Reinelt, 1985, qui traite du problème plus général évoqué ci-dessus).

Terminons cet exposé par deux remarques. Il est intéressant de constater que les trois principales redécouvertes de la règle de Condorcet ont eu lieu pratiquement au même moment (1959-60), mais qu'une seule a été largement connue, celle de Kemeny, certainement à cause de l'ouvrage de Kemeny et Snell ; on vérifie ainsi, une fois de plus, qu'une bonne idée connaît souvent plusieurs naissances indépendantes lorsqu'elle est mûre, mais aussi qu'il ne suffit pas qu'elle soit bonne pour qu'elle prospère. On a vu plus haut que le problème d'optimisation de Condorcet était un cas particulier d'un problème d'optimisation combinatoire très général, mais consistant toujours dans la recherche d'un ordre ; dans une autre direction la notion d'ordre médian peut s'étendre à beaucoup d'autres types de relations, et est notamment intéressante en classification dans la recherche d'équivalences médianes ; on pourra trouver une présentation des travaux - nombreux - menés sur la "procédure médiane" dans Barthélemy et Monjardet (1981, 1988).

RÉFÉRENCES

ADELSMAN, R., WHINSTON, A., The equivalence of three social decision functions, *Revue d'Automatique, Informatique et Recherche Opérationnelle* 11 (3) (1973) 257-265.

ARROW, K.A., RAYNAUD, H., *Social Choice and multicriteria decision-making*, Cambridge, M.I.T. Press, 1986.

BARBUT, M., Médiane, distributivité, éloignements, *Publications du Centre de Mathématique Sociale*, E.P.H.E. 6e section, Paris (1961) et *Math. Sci. hum.* 70 (1980) 5-31.

BARBUT, M., Note sur les ordres totaux à distance minimum d'une relation binaire donnée, *Math. Sci. hum.*17 (1966), 47-48.

BARBUT, M., Médianes, Condorcet et Kendall, *note SEMA*, Paris, 1967, et *Math. Sci. hum.* 69 (1980) 5-13.

BARBUT, M. , MONJARDET, B., *Ordre et Classification, Algèbre et Combinatoire,* Tomes I et II, Paris, Hachette, 1970.

BARTHÉLEMY, J.P., MONJARDET, B., The median procedure in cluster analysis and social choice theory, *Math. Soc. Sci.* 1 (1981) 1 235-268.

BARTHÉLEMY, J.P., MONJARDET, B.,The median procedure in data analysis : new results and open problems, in *Classification and related methods of data analysis*, H.H. Bock ed., Elsevier, 1988, 309-316.

BLIN, J.M., WHINSTON, A.B., Combinatorial optimization and preference pattern aggregation, *Lectures Notes in Computer Science*, Berlin, Springer Verlag, 1972, 73-84.

BLIN, J.M., WHINSTON, A.B., Discriminant functions and majority voting, *Manag. Sci.* 21 (1975), 1029-1041.

BOWMAN, V.J. , COLANTONI, C.S., Majority rule under transitivity constraints, *Manag. Sci.* 19 (1973), 1029-1041.

BOWMAN, V.J., COLANTONI, C.S., Further comments on majority rule under transitivity constraints, *Manag. Sci.* 20 (1974) 1441.

BRUNK, H.O., Mathematical models for ranking from paired comparisons, *J. Amer. Statist. Assoc.* 55 (1960) 503-520.

Marquis de CONDORCET M.J.A ., *Essai sur l'application de l'analyse à la probabilité des décisions rendues à la pluralité des voix* (Paris, 1785) (reprint, New York, Chelsea Publ., 1974).

CRÉPEL, P., Le dernier mot de Condorcet sur les élections, *Math. Inf.et Sci. hum.*111 (1990) 7-43.

De CANI, J.S., Maximum likelihood paired comparison ranking by linear programming, *Biometrika* 56 (3) (1969) 537-545.

DEGENNE, A., *Techniques Ordinales en Analyse des Données : Statistique*, Paris, Hachette, 1972.

FISHBURN, P.C., Condorcet social choice function, SIAM, *J. Appl. Math.* 33 (3) (1977) 469-489.

GUÉNOCHE, A., Un algorithme pour pallier l'effet Condorcet, *Revue d'Automatique, Informatique et Recherche Opérationnelle* 11 (1) (1977) 77-83.

GUILBAUD, G. Th., Les théories de l'intérêt général et le problème logique de l'agrégation, *Economie Appliquée* 5 (4) (1952), reprinted in Eléments de la théorie des Jeux (Paris, Dunod, 1968), English Translation in Readings in *Mathematical Social Sciences* (Science Research Associates, Chicago, 1966) 262-307.

HAYS, W.L., A note on average τ as measure of concordance, *J. Amer. Statist. Assoc.* 55 (290) (1960) 331-341.

JACQUET-LAGRÈZE, E., L'agrégation des opinions individuelles, *Informatique et Sciences humaines* 4 (1969) 1-21,

JACQUET-LAGRÈZE, E., Analyse d'opinions valuées et graphes de préférences, *Math. Sci. hum.* 33 (1971) 33-55.

KEMENY, J.G., Mathematics without numbers, *Daedalus* 88 (1959) 577-591.

KEMENY, J.G., SNELL J.C., *Mathematical Models in the Social Sciences* (Ginand Co, New York, 1961).

KEMENY, J.G., *Rank Correlation Methods* (Hafner, New York, 1962) 3rd edition.

KENDALL, M.G., Note on the estimation of a ranking, *J. R. Statis. Soc.*, 105, 1942, 119.

LEVENGLICK, A., Fair and reasonable election systems, *Behavioral Science* 20, 1975, 34-46.

MARCOTORCHINO, J.F. , MICHAUD, J.P., *Optimisation en Analyse Ordinale des Données*, (Masson, Paris, 1979).

MERCHANT D.K., RAO, R., Majority decisions and transitivity : some special cases, *Manag. Sci.* 23 (2) (1976) 12130.

MICHAUD, J.P., Agrégation à la majorité ; hommage à Condorcet, Centre Scientifique IBM France, *Etudes* n° F. 0 51 (1982) et F. O 94 (1985)

MIRKIN, B.G., The problems of approximation in space of relations and qualitative data analysis, *Automatika i Telemechanica*, translated in *Automation and Remote Control* (Approximation problems in a relation space and the analysis of non numeric methods) 35 (9) (1974)) 1424-1431.

MONJARDET, B., Tournois et ordres médians, *Math. Sci. hum.* 43 (1973) 55-70.

MONJARDET, B., Concordance et consensus d'ordres totaux : les coefficients K et W., *Revue de Statistique Appliquée* 33, 2 (1985) 55-87.

MONJARDET, B., "Éléments pour une histoire de la médiane métrique", *Documents CAMS*, P056, 1990 ; à paraître dans *Moyenne, milieu et centre : histoires et usages*, J. Feldman, G. Lagneau, B. Matalon, édit., Paris, Éditions de l'E.H.E.S.S., 1991.

REINELT, G., *The linear ordering problem : algorithms and applications*, Berlin, Heldermann Verlag, 1985.

YOUNG, H.P., "Condorcet's theory of voting", *American Political Science Review*, 82 (4) (1988), réimprimé in *Math. Inf et Sci. hum.*, 111 (1990) 45-59.

YOUNG, H.P., LEVENGLICK, A., A consistent extension of Condorcet's election principle, *SIAM, J. Appl. Math.* 35 (2); (1978) 285-300.